

BAB II

METODOLOGI PENELITIAN

2.1 Obyek Penelitian

Obyek pada penelitian ini adalah menganalisis sentimen masyarakat terhadap adanya Jembatan Repo-Repo. Mengapa objek penelitian ini harus di angkat karena dapat memberikan pandangan masyarakat. Dengan mengidentifikasi harapan, kekhawatiran, dan dukungan, sehingga dapat merumuskan strategi komunikasi dan kebijakan yang lebih efektif dan efisien terhadap masyarakat yang berkunjung.

Penelitian mengenai sentimen terhadap bangunan Jembatan Repo-Repo di Google Maps memiliki dampak dalam memahami pandangan masyarakat yang berkunjung dan bisa mengambil keputusan yang lebih baik. Hasil dari analisis sentimen dapat digunakan untuk merumuskan kebijakan yang lebih responsif. Dengan mengetahui sentimen masyarakat, pemerintah dapat mengambil keputusan yang lebih bijak dan diterima oleh masyarakat yang berkunjung.

Penelitian ini dapat mendorong partisipasi masyarakat yang lebih besar dengan menunjukkan pendapat dan perasaan mereka diperhatikan. Ini dapat meningkatkan rasa memiliki dan memelihara terhadap bangunan Jembatan Repo-Repo yang ada di Kutai Kartanegara. Agar masyarakat yang dari dalam negeri maupun luar negeri akan mendapatkan kepuasan dan kesenangan saat melewati Jembatan.

Adapun dalam penelitian ini aplikasi sistem informasi yang digunakan dalam mengambil data sentimen merupakan aplikasi Google Maps. Mengapa tidak menggunakan aplikasi lain atau platform media, karena aplikasi Google Maps mempunyai beberapa keunggulan dalam pengambilan data sentimen yaitu:

- (i) Sumber Data Penting: Ulasan dapat dijadikan sumber data penting untuk penelitian Neuro-Linguistic Programming(NLP), termasuk analisis sentimen.
- (ii) Akses Data: Google Maps adalah salah satu aplikasi yang sangat berguna untuk melakukan pengumpulan data guna menganalisis sentimen masyarakat.
- (iii) Penggunaan Algoritma: Setelah mengumpulkan data, Google Maps memungkinkan algoritma Naive Bayes untuk analisis sentimen masyarakat.

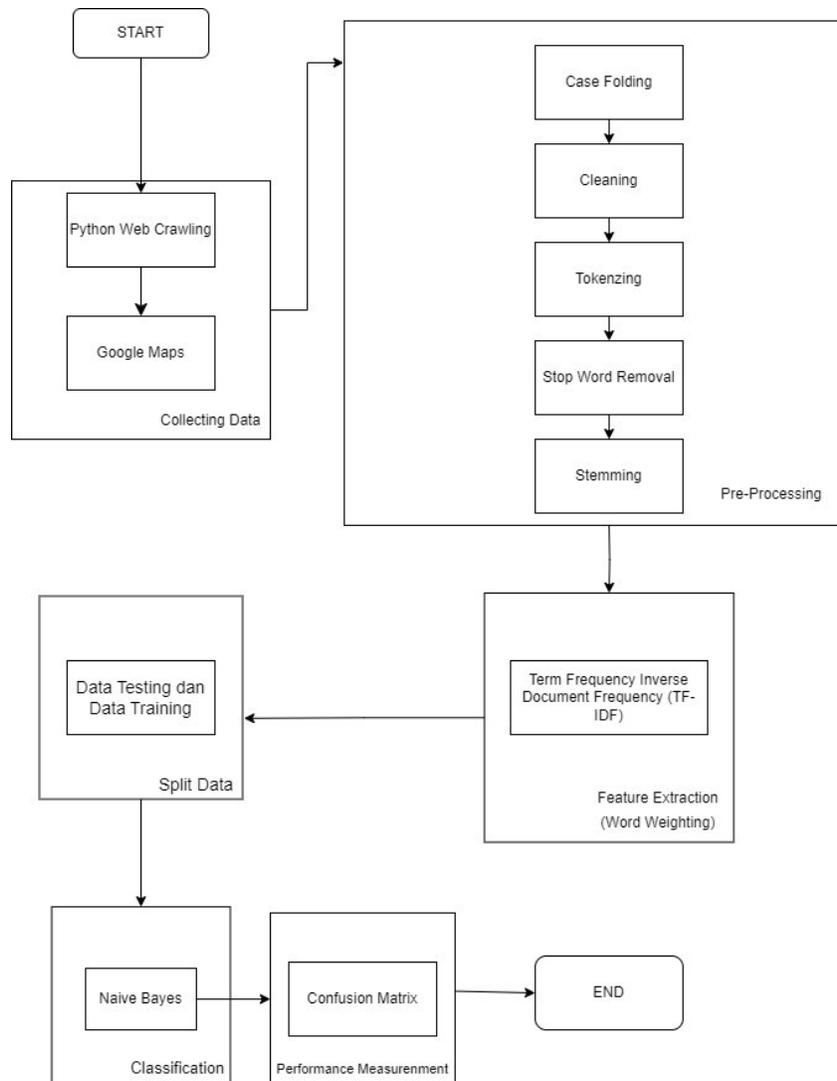
2.2 Alat dan Bahan

Alat yang harus disiapkan pada penelitian ini berupa tools yang telah disediakan Google yaitu Google Colab atau Google Colaboratory. Salah satu tools pengembangan machine learning yang dinilai user-friendly. Google Colab adalah fasilitas bahasa pemrograman Python dari Google, fasilitas ini memungkinkan para pengguna Google Colab untuk saling berkolaborasi dalam membuat pemrograman dengan menggunakan python.

Bahan yang harus dipersiapkan yaitu berupa data Google Maps, Label Sentimen, Metode Analisis yaitu penelitian ini menggunakan metode Naive Bayes. Kemudian melakukan Preprocessing Data menggunakan tools Google Colab, melakukan tahap pembobotan kata. Setelah tahap-tahap tadi sudah dilakukan peneliti melakukan Evaluasi Model seperti Confusion Matrix untuk mengukur kinerja model yang dilatih.

2.3 Prosedur Penelitian

Dalam penelitian ini, dilakukan teknik web scraping pada Google Maps menggunakan bahasa pemrograman Python kemudian mengambil link Jembatan Repo-Repo pada Google Maps untuk proses analisis lebih lanjut. Metode yang dipakai dalam penelitian ini bertujuan untuk melakukan analisis sentimen terhadap ulasan pada Google Maps dengan menggunakan metode Naive Bayes. Diagram alur penelitian dapat dilihat pada Gambar 2.1



Gambar 2. 1 Alur Penelitian

Pada penelitian Metode ini terdiri dari beberapa tahapan utama, mulai dari Pengumpulan data melalui scrapping menggunakan Python, Pre-processing data seperti case folding, cleaning, tokenizing, stop word removal, dan stemming, Pembobotan kata menggunakan metode TF-IDF untuk memberikan nilai pada setiap kata berdasarkan frekuensi dan kepentingannya, split data atau pembagian data menjadi data training dan data testing, klasifikasi menggunakan metode Naive Bayes, serta Evaluasi kinerja model menggunakan confusion matriks. Setiap tahapan akan dijelaskan secara rinci dengan memaparkan teori, konsep, dan algoritma yang digunakan.

2.3.1 Pengumpulan Data

Dalam penelitian ini pengambilan data dilakukan melalui proses crawling data ulasan Google Maps dengan memanfaatkan library BeautifulSoup4 yang diimplementasikan dalam bahasa pemrograman Python. BeautifulSoup4 digunakan sebagai alat untuk mengumpulkan data ulasan dari Google berdasarkan identifikasi unik sebuah tempat (*feature_id*) dan menganalisis informasi yang terkandung dalam ulasan tersebut, seperti rating dan sentimen pengguna. Dengan menggunakan BeautifulSoup4, peneliti dapat mengekstrak data ulasan yang relevan, mengidentifikasi sentimen dari ulasan (positif, netral atau negatif), dan mengumpulkan informasi penting lainnya dari ulasan pengguna (Ulfah & Najiah, 2023). Hasil crawling akan ditampilkan pada Gambar 2.2

	name	rating	review	time	sentimen
0	Erik Nawaca Harahap	5	Bagus jembatannya tapi tutup dimalam hari, tap...	31-03-2024	Positif
1	firmansyah zeppy	5	Mantapppp,,, rawat dan jaga aset kembangkan u...	01-03-2024	Positif
2	Mita Hady	4	Jembatan penyebrangan ke pulau kumala. Pas des...	31-05-2023	Positif
3	Ros Aizam	5	Pas kesini lg ada pengecatan bagus banget liat...	02-12-2023	Positif
4	Supri Yanus	5	Pernah kesini pada 7 Maret 2019...Bisa berjala...	31-05-2022	Positif

Gambar 2. 2 Hasil Pengumpulan Data

2.3.2 Pre-Processing

Preprocessing menjadi tahap awal dalam klasifikasi teks untuk mempersiapkan data teks sebelum digunakan pada proses lainnya. Pada tahap ini akan mengubah data teks menjadi bentuk yang lebih baik sehingga menghasilkan informasi teks dengan kualitas yang baik dan siap digunakan pada proses selanjutnya (Khairunnisa et al., 2021). Berikut langkah-langkah yang dilakukan pada tahap preprocessing yaitu (Puspitasari et al., 2023) :

- (i) *Case folding* merupakan tahapan untuk mengubah huruf kapital dalam dokumen menjadi standar, yaitu huruf kecil (lowercase).

- (ii) *Cleaning* merupakan tahapan untuk membersihkan karakter yang tidak dibutuhkan seperti URL, @, #, https: RT (Retweet), angka, simbol, dan emoticon.
- (iii) *Tokenizing* merupakan tahapan untuk memecah kalimat dalam dokumen menjadi kata. Sebuah proses untuk menghilangkan tanda baca, simbol, dan karakter bacaan yang tidak berharga.
- (iv) *Stopword removal* merupakan tahapan untuk menghilangkan sebuah kata informatif rendah. Stopword dilakukan jika kalimat tersebut mengandung kata umum dan tidak signifikan seperti waktu, penghubung, dan sebagainya.
- (v) *Stemming* merupakan tahapan untuk menghapus imbuhan awalan dan akhiran, sehingga menjadi bentuk kata dasar. Stemming dilakukan menggunakan library Sastrawi pada python.

2.3.3 Pembobotan Kata (TF-IDF)

Pembobotan kata bertujuan memberikan bobot pada setiap kata yang terdapat dalam, kalimat/dokumen yang sudah melalui tahap pre-processing. Untuk mendapatkan nilai TF-IDF, akan dilakukan perhitungan Term Frequency(TF), dan Inverse Document Frequency (IDF). Term Frequency (TF) merupakan proses untuk menghitung jumlah frekuensi suatu kata terdapat dalam keseluruhan dokumen dan memberikan nilai untuk kata tersebut (Locarso, 2022).

Penggunaan TF dapat menggunakan rumus di bawah ini:

$$tf_{ij} = \frac{f_d(i)}{\max f_d(j)}$$

TF menunjukkan dokumen (d) seberapa banyak kata (t) yang muncul. Dan terkait untuk rumus IDF yaitu:

$$idf_t = \log \frac{N}{df_t}$$

N melambangkan jumlah kata dalam teks, df adalah jumlah teks yang memiliki kata t. Dengan menggabungkan TF dan IDF dalam pengerjaan dapat membantu meningkatkan performa.

Setelah mendapatkan TF dan IDF, kemudia dapat menghitung nilai TF-IDF yang merupakan hasil perkalian dari TF dan IDF. Contoh rumus yang sering digunakan sebagai berikut.

$$tfidf_{t,d} = tf_{t,d} \times idf_d$$

2.3.4 Split Data

Splitting data adalah metode yang digunakan untuk membagi dataset menjadi data train dan data test. Fungsi dari data train yaitu data yang akan digunakan untuk dilatih pada saat pembuatan model, dan fungsi dari data test yaitu untuk data yang akan diuji dari hasil prediksi yang dibuat berdasarkan data train. Pada tahap splitting data, peneliti memutuskan untuk melakukan percobaan dengan membagi data menjadi 90:10, 80:20, dan 70:30 dengan urutan data training dan data testing. (Ariyani et al., 2022).

2.3.5 Klasifikasi

Salah satu algoritma yang digunakan untuk klasifikasi teks adalah Naive Bayes. Algoritma ini menggunakan perhitungan probabilitas dan statistik dalam pengoperasiannya. Metode klasifikasi Naive Bayes digunakan untuk mengambil keputusan dengan melakukan prediksi suatu kasus berdasarkan hasil dari klasifikasi yang telah diperoleh. Pada penelitian ini, metode Naive Bayes digunakan untuk menentukan sentimen dari data Google Maps (Astari et al., 2020). Berikut ini adalah langkah – langkah rumus Naïve Bayes yaitu:

$$P(A|B) = \frac{P(B|A) \times P(A)}{P(B)}$$

Di mana:

A: Hipotesis data merupakan suatu *class* spesifik.

B: Data dengan kelas yang masih belum diketahui.

$P(A|B)$: Probabilitas hipotesis berdasar kondisi.

$P(A)$: Probabilitas hipotesis.

$P(B|A)$: Probabilitas berdasarkan kondisi pada hipotesis.

$P(B)$: Probabilitas B.

2.3.6 Evaluasi

Evaluasi digunakan untuk melihat kinerja dari model klasifikasi yang digunakan. Metode evaluasi yang digunakan adalah confusion matrix. Berdasarkan confusion matrix pengujian model dilakukan dengan menghitung nilai accuracy, precision, recall, dan juga F-measure (Muslimin & Lusiana, 2023).

Perhitungan untuk metode confusion matrix yang digunakan yaitu nilai skor Accuracy, karena untuk mengukur seberapa sering model klasifikasi membuat prediksi yang benar, rumus berikut digunakan untuk menghitung nilai akurasi:

$$Akurasi = \frac{TP + TN}{TP + TN + FP + FN}$$

Keterangan:

(i) TP = True Positive

(ii) FN = False Negative

(iii) FP = False Positive

(iv) TN = True Negative

2.3.7 Jadwal Penelitian

Tabel 2. 1 Jadwal Penelitian

No	Kegiatan	Bulan/2024					
		Februari	Maret	April	Mei	Juni	Juli
Tahap Pra Penelitian							
1	Menentukan Judul						
2	Identifikasi Masalah						
3	Studi Literatur						
4	Rancangan Metode						
5	Pemilihan Studi Kasus						
6	Menyusun Proposal						
Tahap Penelitian							
1	Pengumpulan Data (Crawling)						
2	Labelling Data						
3	Pre-Processing Data						
4	Pembobotan Kata (TF-IDF)						
5	Split Data						
6	Klasifikasi Support Vector Machine (SVM)						
7	Evaluasi dan Analisis Hasil						
Tahap Akhir Penelitian							
1	Penyusunan Laporan						
2	Seminar Hasil						